NAWCWPNS TP 8348

Partial Differential Equations for Image Processing

by
Dr. Gary Hewer
Guidance and Control Systems Division
and
Dr. Richard Lau
Mathematical, Computer, and Information Sciences Division
Office of Naval Research

with contributions by
Professor Nathan Intrator and Dr. Quyen Huynh, Brown University
Dr. Leonid Rudin, Cognitech, Inc.
Dr. Guillermo Sapiro, Hewlett-Packard Laboratories
Professor Rene Carmona, Princeton University
Professor James Sethian, University of California, Berkeley
Professor Tony Chan, University of California, Los Angeles
Professor Stanley Osher, University of California, Los Angeles
Professor Bangalore Manjunath and Dr. Charles Kenney, UC Santa Barbara

MARCH 1997

NAVAL AIR WARFARE CENTER WEAPONS DIVISION
CHINA LAKE, CA 93555-6100

Naval Air Warfare Center Weapons Division

FOREWORD

This report, prepared for the Rapid Retargeting Accelerated Capabilities Initiative under the sponsorship of the Office of Naval Research, provides a high-level description of how partial differential variational techniques are applied to image processing. The overall scientific theme is modeling image processing using the versatile mathematics of variational methods solved by partial differential equations (PDEs). This model is new within the last decade and generates both new and superior methods and useful generalizations of known methods. These generalizations allow both better insight into the methods and better performance. A case can be made that these generalizations have sparked a renewal in the theory—borrowing from computational fluid dynamics and differential geometry—and numerics of a class of nonlinear PDEs. The report gives selected examples of the major types of investigation going on and illustrates them with sample imagery. The report also describes how these investigations can be connected to Navy needs in image exploitation and in template-based assisted or automatic target recognition. The challenge of this accelerated capabilities initiative is to identify important applications or opportunities within the rapid retargeting sensor-to-shooter structure that can be improved or developed using these methods.

This report was reviewed for technical accuracy by Dr. Charles D. Creusere, NAWCWPNS Code 4B4000D.

Approved by
A. MICKELSEN, *Acting Head*
Guidance and Control Systems Division
5 March 1997

Under authority of
J. V. CHENEVEY
RADM, U.S. Navy
Commander

Released for publication by
S. HAALAND
Director for Research and Engineering

NAWCWPNS Technical Publication 8348

Published by .... Technical Information Division
Collation .... Cover, 19 leaves
First printing .... 30 copies

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1997	Final, fiscal year 1996

4. TITLE AND SUBTITLE Partial Differential Equations for Image Processing	5. FUNDING NUMBERS
6. AUTHOR(S) Gary Hewer and Richard Lau	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Commander (Code 472C00D) Naval Air Warfare Center Weapons Division 1 Administration Circle China Lake, CA 93555-6100	8. PERFORMING ORGANIZATION REPORT NUMBER NAWCWPNS TP 8348
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Naval Research 800 North Quincy Street Arlington, VA 22217-5660 Ballston Tower 1	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES	
12A. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.	12B. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

(U) This report, prepared for the Rapid Retargeting Accelerated Capabilities Initiative, provides a high-level description of how partial differential variational techniques are applied to image processing. The overall scientific theme is modeling image processing using the versatile mathematics of variational methods solved by partial differential equations (PDEs). This model is new within the last decade and generates both new and superior methods and useful generalizations of known methods. These generalizations allow both better insight into the methods and better performance. A case can be made that these generalizations have sparked a renewal in the theory—borrowing from computational fluid dynamics and differential geometry—and numerics of a class of nonlinear PDEs. The report gives selected examples of the major types of investigation going on and illustrates them with sample imagery. The report also describes how these investigations can be connected to Navy needs in image exploitation and in template-based assisted or automatic target recognition. The challenge of this accelerated capabilities initiative is to identify important applications or opportunities within the rapid retargeting sensor-to-shooter structure that can be improved or developed using these methods.

14. SUBJECT TERMS	15. NUMBER OF PAGES 35
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18
298-102

CONTENTS

INTRODUCTION

This report provides a high-level description of how partial differential variational techniques are applied to image processing. The overall scientific theme is modeling image processing using the versatile mathematics of variational methods solved by partial differential equations (PDEs). This model is new within the last decade and generates both new and superior methods and useful generalizations of known methods. These generalizations allow both better insight into the methods and better performance. A case can be made that these generalizations have sparked a renewal in the theory—borrowing from computational fluid dynamics and differential geometry—and numerics of a class of nonlinear PDEs. The report gives selected examples of the major types of investigation going on and illustrates them with sample imagery. The report also describes how these investigations can be connected to Navy needs in image exploitation and in template-based assisted or automatic target recognition. The challenge of this accelerated capabilities initiative is to identify important applications or opportunities within the rapid retargeting sensor-to-shooter structure that can be improved or developed using these methods.

This report has multiple aims: 1) to introduce a new and more effective image-processing technology, 2) to provide a status report on progress in a new Office of Naval Research (ONR) basic research (6.1) program in the area, and 3) to discuss the relevance of this 6.1 program to development issues (6.2-6.3). Development issues are discussed in the section Navy Needs and Background, with emphasis on rapid retargeting; work aimed at sea-mine identification is largely suppressed here. We expect the reader to be able to map the research results and issues of the section PDE-Based Multiscale Image and Video Processing onto the issues presented in the previous section. The authors believe that they have, under the aegis of the Rapid Retargeting accelerated capability initiative (ACI), implemented an integrated 6.1-6.2 program, but to point out all of the detailed connections and interactions would be tedious.

The basic research program is supporting research into variational methods in image processing. These techniques lead to fundamental research into the application of a wide set of algorithms, whose core is based on PDEs as a model for image processing. The PDE appears as one way to approximate the minimizing solution of the variational problem. The 6.2 thread of this program involves sensor-specific investigations to evaluate the core algorithms and identify transition paths and applications in the sensor-to-shooter chain that will ultimately enable real-time retargeting capability for future weapon systems and tactical strike manned aircraft. The 6.2 research coordinates with the 6.3a Cruise Missile Real-Time Retargeting Demonstration Program (CMRTR Demo) by focusing on the laser radar (LADAR) mobile target automatic target recognition (ATR) system.

Whereas 6.1, 6.2, and 6.3 do, to a degree, "live in different time zones," a thesis of this program is that they can cooperate significantly and to their considerable mutual advantage. The most important reason that cooperation or integration is workable here is that the R&D product is algorithms, which are not terribly hard to implement and can be tried out on practical problems very quickly. Further, testing against real problems provides important feedback to the 6.1 methods developers. Another important aid to integration is that many of the PDE methods both improve and generalize traditional image-processing methods, thereby lowering the transition hurdles. Finally, PDE methods provide a fundamentally new paradigm for the study of images—one that yields demonstrably better methodologies. This paradigm produces better mousetraps in addition to better science. Partial differential equations have provided the mathematical backbone of physics since Newton. A very rich body of mathematics has been generated for their study. A case can be made that we are taking image processing out of the realm of heuristics and into the territory of physics. The research will replace "cut and try" methods with a mathematical theory so that the subject can be and should be viewed as image science and not just as tinkering. The theory should fit image processing as well as calculus fits classical physics.

The computer bridge between 6.2 and 6.3 is often provided by a system simulation, which can be used for rapid prototyping. Practical constraints caused by 6.3 hardware configurations, project schedules, and real-time processing requirements imposed by missile platforms and sensors are being addressed when assessing opportunities for algorithm transition, as are performance metrics (e.g., what contrast ratio is needed to support the partial derivative estimates in the PDE arena).

The next section describes the application threads to the military problems. It identifies the functional context for two transition paths, namely, image exploitation aids for rapid retargeting and model-based ATR. The following section sketches the overall scientific theme-modeling image processing using the mathematics of partial differential equations. Both sections contain a mosaic of imagery highlighting some of the research accomplishments. Subsequently, a brief description of the major research investments is provided.

NAVY NEEDS AND BACKGROUND

Navy Needs

An important lesson from the Gulf War was that the information age is arriving in the battle space. Words like comprehensive awareness, dynamic replanning, and dominant maneuver are often used to describe future tactical scenarios. A key area is targeting intelligence. An important component of targeting intelligence is networking the existing reconnaissance databases within the time cycles dictated by the needs of rapid retargeting, which is also the theme of this ACI. A brief attack window is one of the quintessential tactical descriptions associated with the precision attack of fixed and rapidly relocatable mobile targets (e.g., Scud missile launchers). To exploit this window requires real-time coordination among the surveillance, targeting, battle damage assessment, and mission planning systems that support the mission. Each of these network components depends on real-time sensor information from a variety of platforms (e.g., satellites and aircraft) and sensors (e.g., multispectral infrared images, synthetic aperture radar (SAR) images, and digital terrain maps), which must be fused with prior information for dynamic replanning. The cartoon in Figure 1 suggests how the networks are formed.

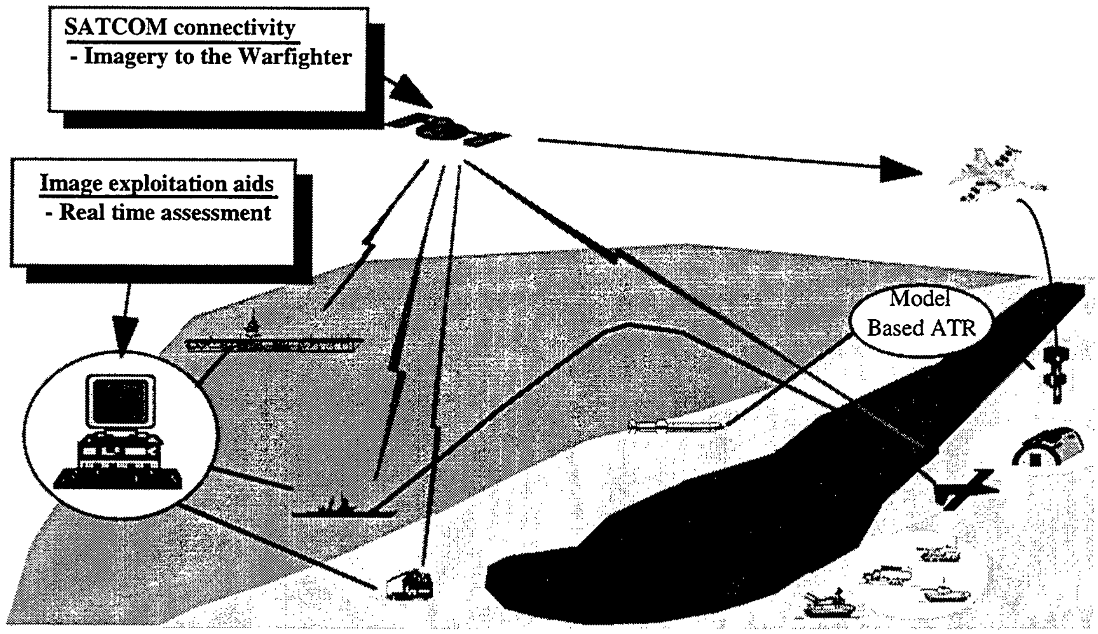


FIGURE 1. Sensor-to-Shooter Links.

The two key integration nodes identified for the ACI image-processing applications are 1) any workstation node requiring image exploitation aids, and 2) any autonomous weapon node employing model-based ATR. These nodes are gateways that define the transition threads for the integrated research plan. Our approach for both nodes is a common theoretical formalism that will guide the integration and automation of the key image-processing algorithms required from sensor to shooter.

Image Exploitation

Images (e.g., photo, video, and IR) contain a great deal of information that is not accessible to a human observer no matter how good his observational tools. Images are also degraded by environmental factors, copying, etc. The images in Figure 2 illustrate this hidden information and the degree to which PDE methods are effective. The original image is a video of a crime scene, taken from a helicopter. One 6000th of it was of interest in the Los Angeles riots trial. The left image in the second row is a high-tech enlargement of the perpetrator's left arm. The other images in the row were segmented (variational) by PDE methods. The middle image shows the segmented arm, just delineating the boundaries, and the right image is the full piecewise constant segmentation, which shows a clear outline of the rose tattoo on the arm. The bottom row shows the dramatic agreement with the actual tattoo. The suspect was identified and convicted with the help of this evidence.

The images in Figure 2 dramatically illustrate the extent to which advanced enhancement and restoration methods can both restore degraded images and uncover hidden information. Functionally, most methods lie within the following image-processing taxonomy, namely, contrast enhancement, restoration, segmentation, fusion, registration, browsing, mining, and videogrammetry. Contrast enhancement is a basic and important operation, which improves the image quality for the viewer by modifying the gray scale. Restoration techniques include fusion, and removal of noise and blur. Segmentation is basically a process of pixel classification into classes that can include dark and light or, more commonly, edges and non-edge. Image fusion establishes a match or correspondence between several images or between one image and a model. Video browsing and mining are suggestive concepts whose formulation and potential are yet undeveloped. Examples within this class include intraframe and interframe feature extraction (e.g., parking lots, buildings, and landmarks), morphological typing (e.g., roof tops), and interframe change detection using velocity (optical) flow field estimation. Feature extraction can be used for recognition or to suppress redundant information. Registration overlays grids such as earth coordinates so that topographic features such as roads or buildings can be measured.

One of the central tasks of photogrammetry is the quantitative localization and reconstruction of objects in the scene for measurement or interpretation—applications include targeting or battle damage assessment. Videogrammetry is a recent extension of photogrammetry to the digital domain. Although present image-processing texts follow this taxonomy, a novelty of this research is to introduce a systematic approach that unifies the principles governing these seemingly separate and unrelated techniques.

Videogrammetry is an extension of the principles of photogrammetry to video imagery. The crash analysis in Figure 3 hints at the undeveloped potential of this new discipline for crash scene measurements.

The image on the left in the first row in Figure 4 is the original low-contrast, blurred image captured by an automatic teller machine (ATM) surveillance camera. Because the primary function of an ATM's cameras is to record ATM patrons, the frame rate is low—only one or two frames per minute—and the focused depth of field is only a few feet. The monitoring role of these cameras makes them serendipitous archives of criminal activity. Because they are not optimally designed for this role, mining this archive is a major challenge. The image on the right in the first row is the restored image, which provided six key identifying features, allowing the actual vehicle to be positively identified with the vehicle in the ATM video. The imagery in the next two rows is self-explanatory.

Presently, image exploitation aids are needed by and designed for image analysts as they execute all of the necessary steps that are required for mission planning, targeting, and battle damage assessment. Performing many of these functions within 10- to 30-minute execution times requires automated processing aids to augment the present labor-intensive "click and point" mouse and cursor interfaces. Thus, our goal for both 6.1 and 6.2 is to either identify those workstation bottlenecks that can be improved, automated, or augmented by the ACI algorithms, or to introduce new exploitation aids tailored to digital libraries.

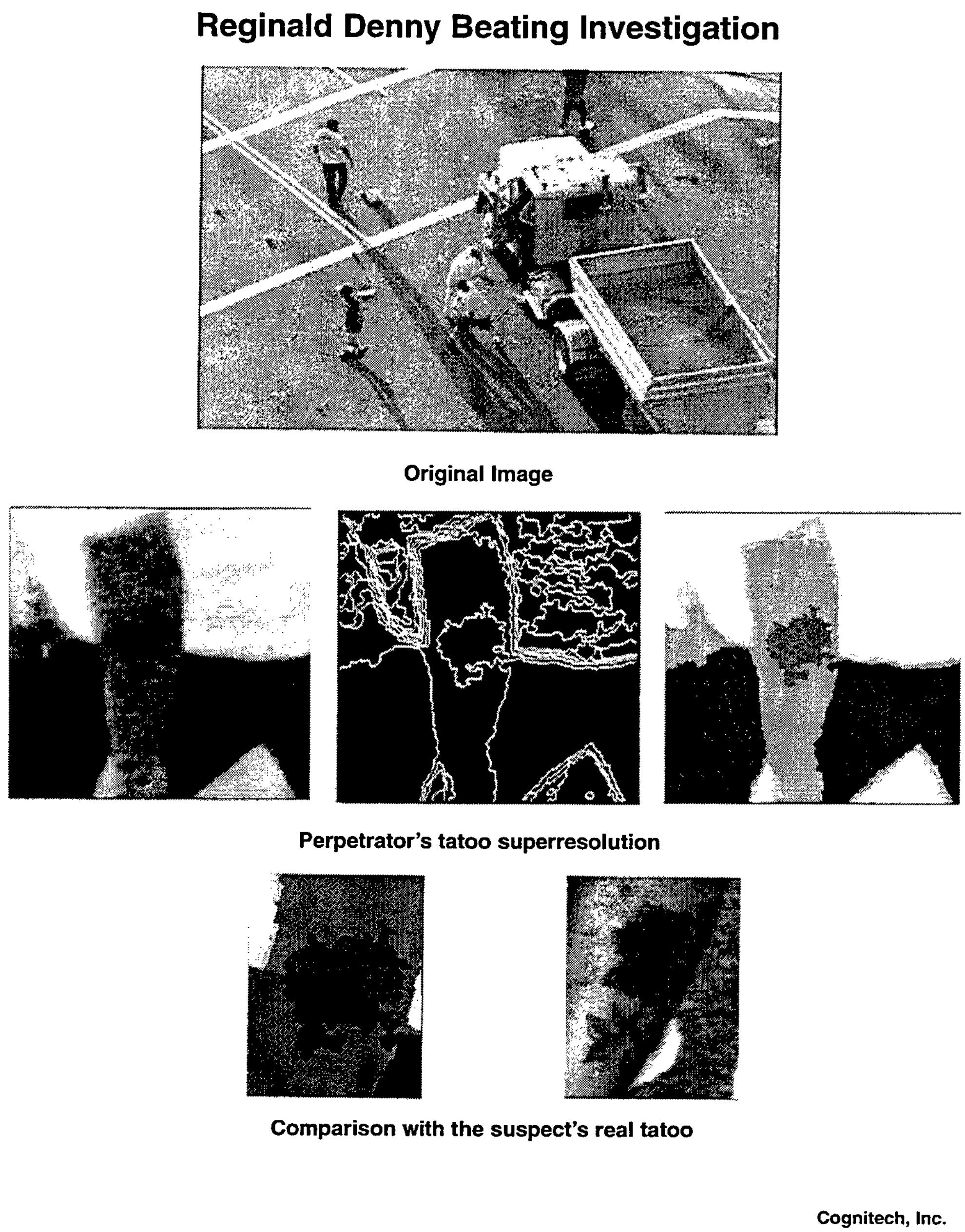
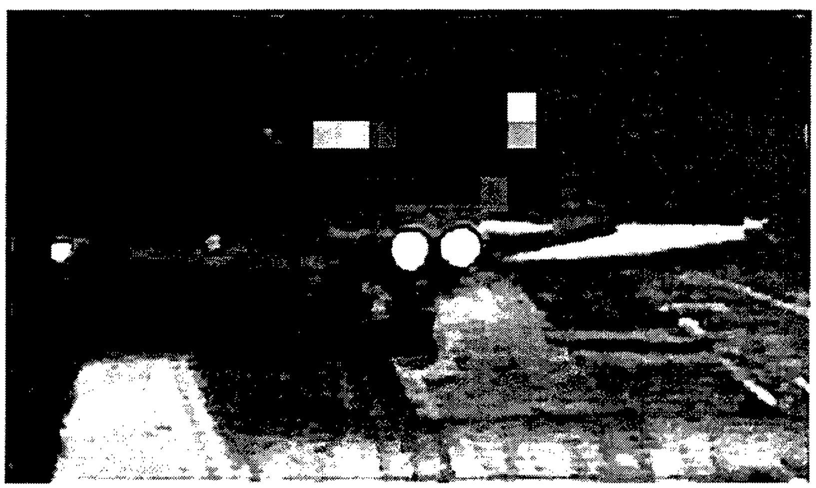


FIGURE 2. Investigative Image Processing.

Naval Aircraft Crash Investigation

Aircraft just after launch, at t=0
Original Image (identification detail removed)

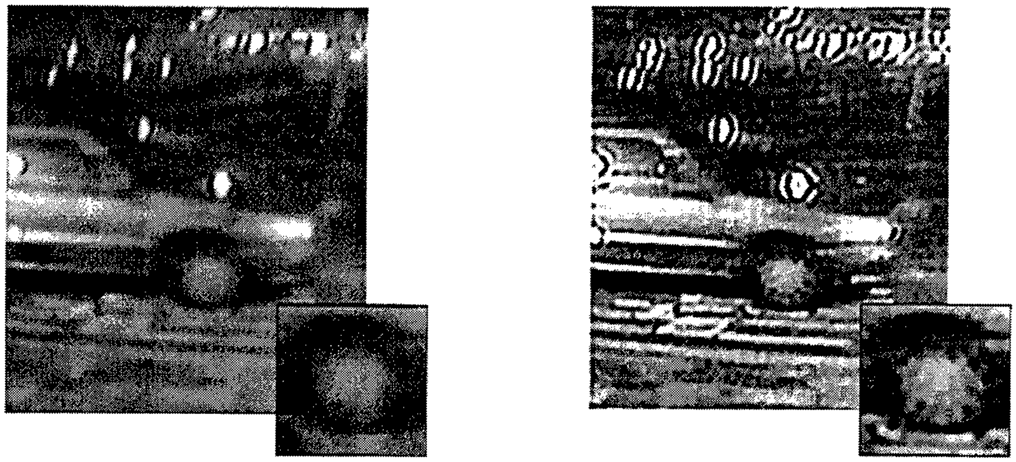


Aircraft just after launch, at t=0.
Image Restored using Cognitech's restoration algorithm
(identification detail removed)

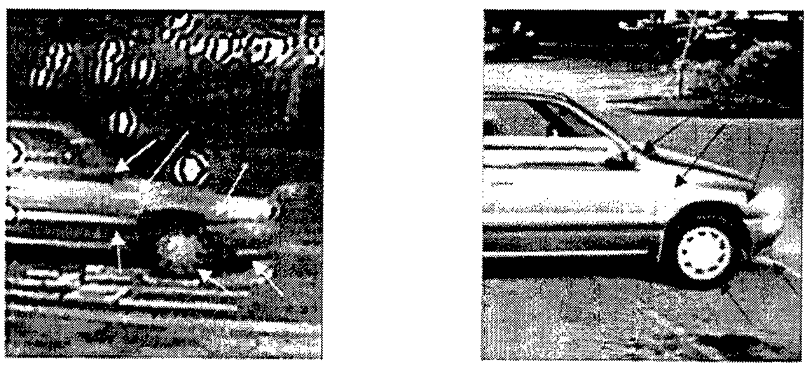
Cognitech, Inc

FIGURE 3. Videogrammetry.

Manhattan Beach Investigation
(Murder of a Police Officer)



Original ATM captured Image **Cognitech's restoration**



Comparing with the suspect's vehicle

Similar unique features (damage) are identified with the suspect's car

Cognitech, Inc.

FIGURE 4. Investigative Videogrammetry.

Template-/Model-Based ATR

Although the human eye can sometimes find objects in noise quite well, machine processing often requires automatic feature detection. Template-based automatic or assisted target recognition systems are complex and are composed of different algorithms applied in sequence. They have three essential components: 1) feature extraction, 2) feature grouping, and 3) interpretation. These are often classified, respectively, as a low-level vision, a mid-level process, and a high-level process, suggesting a progressive organization of the image data designed to elucidate the structure. Low-level feature extraction includes edge detection and corner detection by curvature analysis. Grouping is often accomplished by segmentation. The third stage is often realized by template matching (e.g., template correlation). In each of these processes, tuning parameters (thresholds) are often used. These parameters are traditionally chosen through extensive simulations using a mixture of empirical and synthesized imagery. Accuracy and spurious pixels are major issues affecting overall performance when low-contrast, dense images with partially occluded small objects are processed. For the autonomous weapon node, our 6.1 and 6.2 goal is to identify those image-processing bottlenecks that can be ameliorated by the ACI algorithms. Currently segmentation algorithms have been identified as the most direct link with the CMRTR Demo program, because they perform a critical role in model based ATR. Moreover, recent variational theory has replaced "cut and try" methods, making them a prime example for transition. While none of our imagery includes LADAR, we expect that the variational techniques can be directly applied, because existing LADAR template-based ATR methods require the same low-level features. What is not known, however, is the extent to which the variational techniques can improve the overall performance of template-based ATR systems.

Cruise Missile Real-Time Retargeting Demonstration— An Early Transition Opportunity

In recent years, the Naval Air Warfare Center Weapons Division at China Lake, under ONR sponsorship, has been exploring the potential use of a laser radar for cruise missile upgrades, as well as other weapon applications that require a high-performance imaging seeker. The 6.3a demonstration effort undertaken by ONR to address this issue is titled Cruise Missile Real-Time Retargeting Demonstration. The CMRTR Demo will demonstrate substantial improvement in affordability and performance of strike weapon systems for real-time retargeting. This advanced technology demonstration will include automatic target recognition, which performs in clutter, and responsive replanning, which is an order of magnitude faster than current systems. The 7-year program (FYs 1995-2001) will transition and integrate the Adaptive Mission Control and Sea-Based Mission Planning technologies, developed under 6.2 programs sponsored through the ONR, and the solid-state LADAR, developed under the Air Force Wright Laboratory's Low-Cost Anti-Armor Submunition program, into a flight-testable modular brassboard. This program is designed to address issues and risks associated with automated real-time (re)planning, ATR, battle damage indication, targeting sensors, and all-up guidance system affordability for advanced glide weapon and cruise missile concepts.

PDE-BASED MULTISCALE IMAGE AND VIDEO PROCESSING

The Setting

In the past few years it has become clear that techniques and ideas originating in fields such as computational fluid dynamics (CFD) and other areas of computational physics have an important and direct relevance to image and video processing. So clear, in fact, that Pierre-Louis Lions, one of the recent winners of the Fields medal in mathematics (the first applied mathematician ever to be so honored) has been seriously working in this area for a number of years. In addition, David Mumford, the current head of the International Mathematical Union, himself a former winner of the Fields medal (in Algebraic Geometry) has switched fields to work full-time in this area. The earliest image-processing algorithms utilizing nonlinear PDEs borrowed from methods used in the CFD community in that they were devised to capture discontinuities, which in images occur at edges. Concepts such as morphology, which have been used in the computer vision community, turn out to be directly connected to PDE and scientific computing based ideas—in this case, the recently developed level set method for front propagation. Just as the language of mathematics is known to be unreasonably effective in the physical world in general, the language of nonlinear PDE and multiscale analysis appears to be unreasonably effective in image and video processing and to apply pretty much across the board in image processing.

So what is so great about the so-called PDE approach? It is connected to systematic view of image processing based on a variational formalism that proceeds in two steps, namely, 1) the formulation of a relevant variational problem judiciously chosen, and 2) the efficient search for a solution of this variational problem. Many methods have been proposed. Simulated anealing has been advocated by the Markov Random Field people, but it is long and one never knows if one got to the global minimum! More recently Linear Programming by Interior Point Methods have been proposed by the Princeton investigator Rene Carmona, but this is so far limited to specific problems of TV with linear constraints.

PDEs arise as one way to approach the solution of the variational problem by writing down the Euler-Lagrange equations and interpreting them as a (typically nonlinear) PDE. What is great is that, when this PDE is reasonable (and Lions and his co-workers have worked very hard to tell us when this is the case), then it gives a dynamical system one can iterate until the system relaxes to equilibrium, and this equilibrium is the solution of the variational problem and of the image analysis task. But what is even better is that one initializes the system with the original image and one sees, like a cartoon, the changes in the image that lead to the solution. Even better, a look at the PDE driving this dynamical system gives a very intuitive and appealing interpretation of the steps taken by the dynamical system treating the image. For example, if the PDE is a nonisotropic diffusion equation, then at each step one smoothes the image in the direction of the features of the image and one tries to avoid doing that across the features. This is why one sees the noise disappear and the edges remaining crisp. These PDE algorithms are superior to all the other algorithms, which act like black boxes (give me the input and I spit out the result without you being able to see how I got from one to the other). You can sit at your computer and see the image evolving from one state to another and check that the changes are what you want, and if they are not, you can act on the process. IT IS NOT A BLACK BOX; it is open to changes and fine-tuning. And even if it does not work, at least you can see where things went wrong.

The areas of basic research in image and video processing for which these methods are relevant include image/video restoration (removal of noise and blur), video frame fusion (merging of several poor-quality individual frames from a video into one or more enhanced frames), multiscale segmentation and decluttering, sensor fusion, edge detection, recognition of shapes from shading and/or texture, color processing, and videogrammetry (measurements, useful in crash analysis). These areas are best introduced by displaying processed imagery, and more examples will be given below.

It makes a great deal of difference what mathematical techniques are used for processing such as denoising. It is easiest to see this in one dimension. Panel A of Figure 5 (courtesy of Mary Oman) shows a function u with jumps (edges) and also shows u corrupted by Gaussian noise. We want to reconstruct u given that we know only the noisy version of u. Panels B through F show various reconstruction techniques, of which the wavelet and Fourier techniques will be the most familiar. All methods roughly reconstruct the general shape of u, but the crunch comes with edge reconstruction. Only the total variation (TV) method, discussed below, does a good job of reconstructing edges. The TV approximation is specified mathematically to give a reasonable approximation and to do this with a function that has as few "wiggles" as possible. Exploration of the TV approximation by minimizing the first variation (i.e., Euler-Lagrange) leads directly to the PDE approach to image processing. Alternative formulations using Markov Random Field approaches with the same vartional princples are concurrently being pursued, but will not be discussed here.

When one looks at the digital image from an imaging sensor representing a scene from the natural world, it is impossible to avoid seeing structure. This is universally true whether the digital image was obtained by a imaging sensor operated in the visible, infrared, or laser radar spectral band. With insight from scattering physics, structural patterns can be identified with real objects in synthetic aperture radar images, too. In fact, the goal of variational methods in image processing is to automatically pass from the unstructured digital image represented by a real or complex number at each pixel to some structured features that can be interpreted by an image analyst or by template-based ATR templates in a missile computer.

In theory some structure is always present in the image, and the real question is the effectiveness of the formalism used to extract the structure. While the scattering physics that contributes to that structure will not be discussed, it is not unimportant nor can it be ignored when defining algorithms for specific applications. The image intensity of the original image is described by a real valued function $g(x,y,t)$ representing "intensity," "brightness," or "gray level" at each pixel in a three-dimensional array (grid) with two spatial coordinates x and y and a time coordinate t. The three-dimensional array $g(x,y,t)$ is a video image, while the array $g(x,y)$ with time suppressed is a still-frame image. That is, $g(x,y,t)$ is the magnitude of the image intensity at pixel location (x,y) at each time t. In order to elucidate the structure, it is assumed that basic features such as "edges," gradients of g ($\nabla g = (g_x, g_y)$), the direction perpendicular to the gradient , $\nabla_{\perp} g = (g_y, -g_x)$, and the curvature = $g_{xx} g_y^2 - 2 g_x g_y g_{xy} + g_{yy} g_x^2$ can be estimated from the image. Subscripts refer to partial derivatives; for example g_x , g_y and g_t are the respective partial derivatives of g with respect to x, y and t. The gradient operator, ∇, is defined by

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)$$

Of course, a numerical approximation will be used to compute these quantities at each pixel.

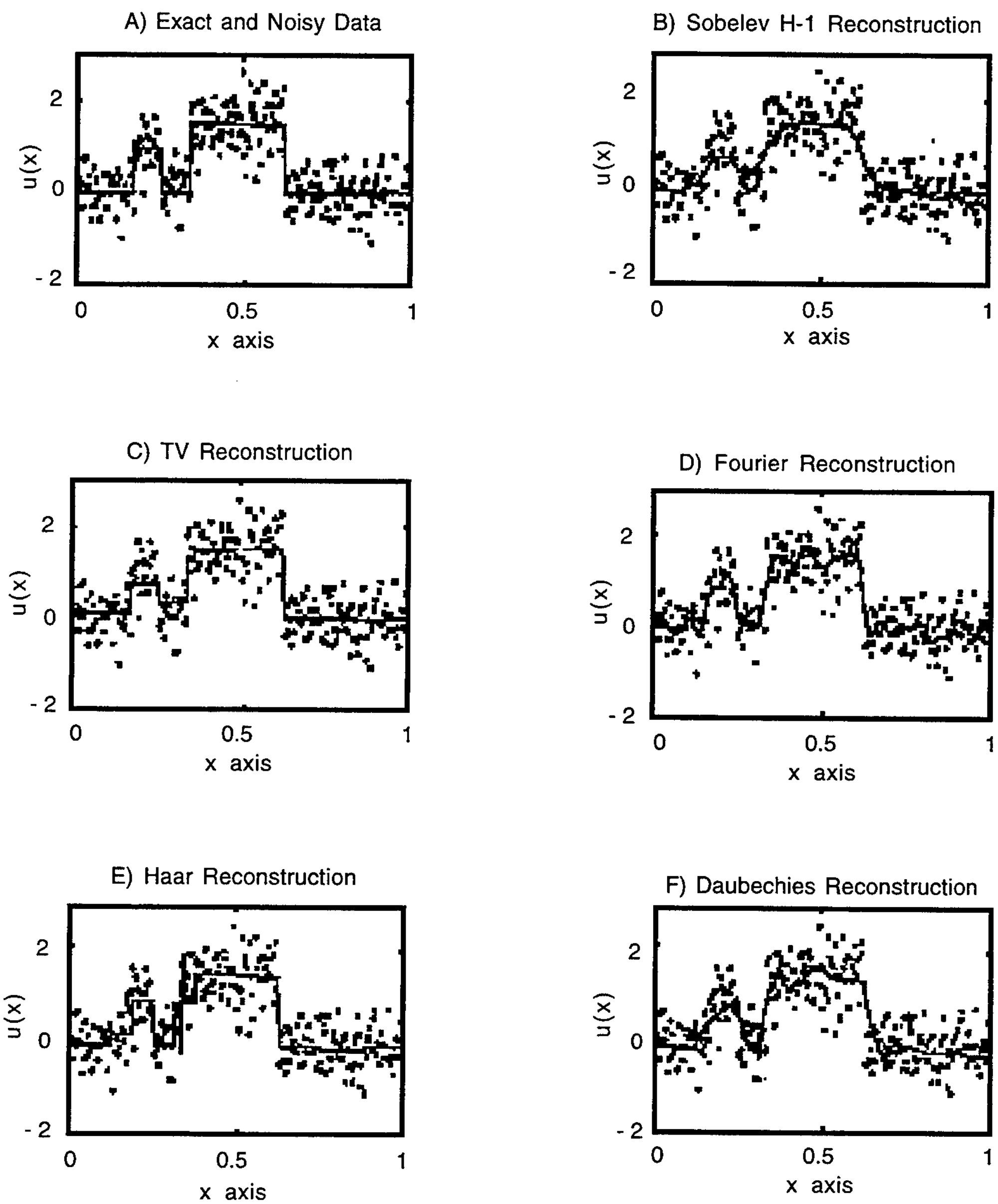


FIGURE 5. Regularization Techniques for One-Dimensional Denoising.

Multiscale Segmentation and Decluttering

The formalism connecting variational methods and partial differential equations will be illustrated by the segmentation problem. Segmenting a digital image means finding homogenous regions R_i and their boundaries B_i. The good contrast in Ansel Adams's black-and-white photographs and in the medium of cartoons makes boundaries identifiable by simply tracing the edges. These special images are not typical of many applications, and more sensitive methods are required. The Mumford-Shah model defines the segmentation problem as a joint smoothing/edge-detection problem. The objective is to compute a segmented image u such that (a) the image u varies smoothly within each region R_i and (b) the segmented image may vary rapidly across boundaries B_i. The objective is achieved by minimizing a weighted objective functional containing three competing terms:

$$E(u,b) = w_1 \int (u - g)^2 \quad + \quad w_2 \int |\nabla u|^2 + \quad w_3 \int_{\sum B_i} db \tag{1}$$

The accuracy of the approximation is driven by the term $\int (u-g)^2$. The smoothness of the approximation is controlled by the term $\int |\nabla u|^2$. The third integral controls the total length of all boundary segments. The motivation behind this approach is that by minimizing the sum of these competing terms we obtain an approximation u that is reasonably close to g, is smooth, and retains the important edge information of g. The 1995 Birkhauser book by Morel and Solimini, *Variational Methods in Image Segmentation*, explains how most known practical segmentation methods translate into the Mumford-Shah model.

Each integral is multiplied by a weight to improve the approximation or the smoothing or to minimize the length of the boundary. By varying the weights, each approximation u results in a different segmented image $u_{seg}(w_1, w_2, w_3)$. Using the functional formulation 3.1 they can be ordered and compared. A criterion for comparing is needed in the segmentation process because the following dilemma arises: 1) edges are the basis of most feature extraction methods, and 2) edges are weak features that are hard to detect reliably. They can be corrupted by noise, blurring, or low contrast. Low contrast occurs when the jump in intensity across an edge is small. Yet edges are the boundaries of image regions.

One of the unifying principles of image analysis is repetitive structure, which leads to hierarchical image structures. The attack pilot's jargon for multiscale structure is funnel features, e.g., progressively detailed salient features that lead to the target. By varying the weights, the Mumford-Shah algorithms decompose the image into a set of multiscale images $u_{seg}(w_1, w_2, w_3)$ indexed by the weights w_i. Thus, the weight parameters determine a scale space decomposition of the original image. This scale space varies from the original image (for the weight choice $w_1 = \infty$) to a completely smooth constant image (for the weight choice $w_2 = \infty$).

We now show an example of multiscale segmentation using three trucks in desert terrain. Figure 6a shows the original image. If a piecewise constant approximation is used, the weights in the Mumford-Shah integral (Equation 1) can be lumped into a single parameter λ. Pure segmentation using three scales is shown in Figures 6b, 6c, and 7a:

$\lambda=40$, $\lambda=32$, and $\lambda=64$, respectively. Note that the trucks are there, but so are the boundaries. The scale parameter λ controls the minimal size of the kept details, causing regions to merge and objects at a specific size to be picked out. An important observation made by L. Rudin is that by inputting various wavelet or other morphological channels into $E(u,b)$ with vector arguments, there exists a remarkable decluttering effect. Figure 6c shows the boundaries of the segmentation of the decluttered image, and Figure 7b shows the reconstruction of the wavelet segmentation. This is all done with no á priori knowledge of what is there. Manmade objects tend to remain and natural texture is dropped.

A major drawback to variational segmentation schemes is that there is no agreed-upon method for selecting the weighting parameters. This limitation does not present a problem in an interactive mode, but it is a limitation for unsupervised segmentation.

Applying the Euler-Lagrange variational procedure to the functional $E(u,b)$ generates a PDE for the approximation u. For the purposes of illustration, letting the weight $w_3 = 0$, we obtain the heat equation

$$u_t = w_1(g - u) + w_2(u_{xx} + u_{yy})$$

As time increases the approximation u evolves in such a way as to decrease the value of the objective functional $E(u,b)$, and solution $u(t,x,y)$ gives the processed image. Usually we define the boundary function in terms of the initial approximation, which satisfies the initial condition $u(x,y,0)=g(x,y)$.

Sensor Fusion

Another important application of these ideas comes when several sensors—e.g., IR, x-ray, visual image, ultrasound, and SAR—view the same object. For our purposes, one simply views the functional Equation 1 as having vector inputs and with appropriate weights; the solution becomes a (nonlinear) superposition of multiple images.

Image Restoration

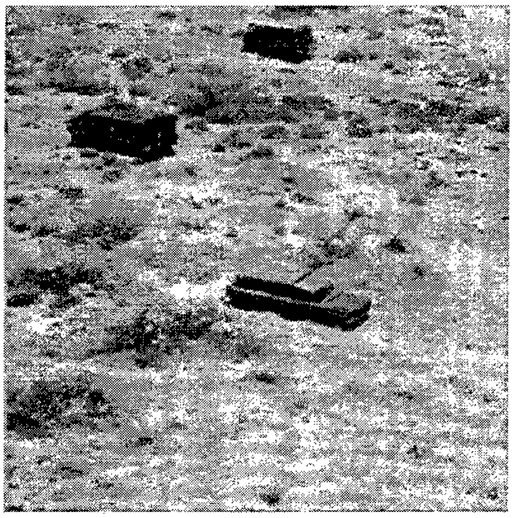
Image restoration is one of the most basic issues in processing. One is presented with an image (or video) that is corrupted by blur (e.g., the picture was taken through a turbulent atmosphere, the camera was moving or out of focus) and noise (familiar to any television viewer). The goal is to restore the original image.

The problem is modeled as follows. Given the blurry, noise image $u_0(x,y)$, where (x,y) are the pixel values in W and u_o is the intensity function, we wish to find $u(x,y)$, the true image. An (oversimplified) model is as follows.

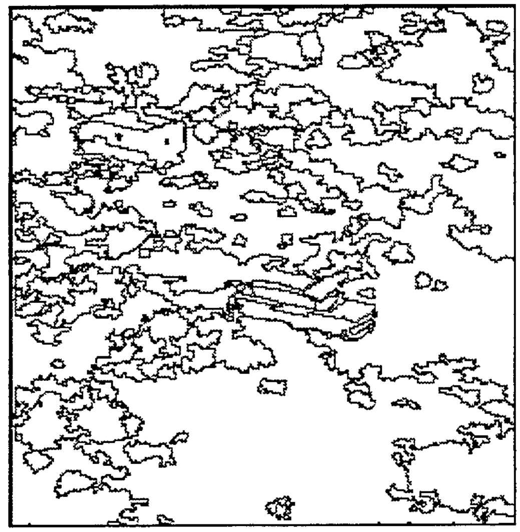
We are given

$$u_0 = Au + n \tag{2}$$

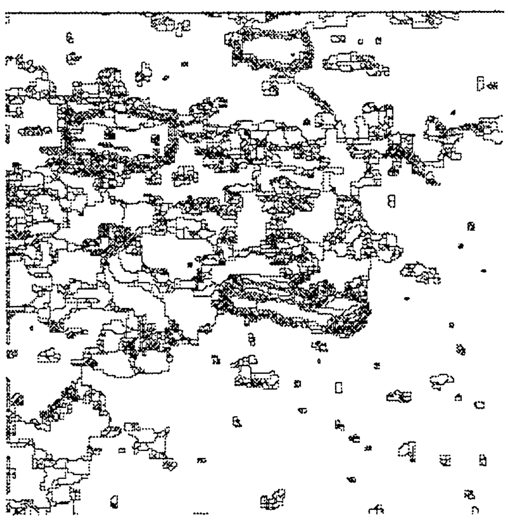
where A is a compact integral operator (e.g., A can represent the sensor optical blur circle) and n is the noise, whose mean and variance can be estimated. Much more complicated models can be dealt with using the new techniques we shall outline.



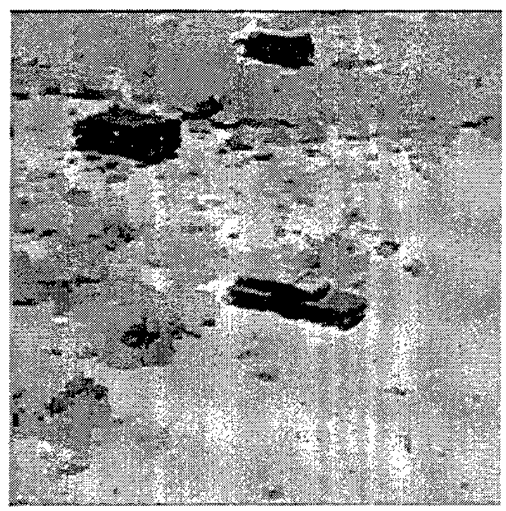
(a) Original picture.



(b)Boundaries of pure gray-level segmentation; $\lambda = 40$.

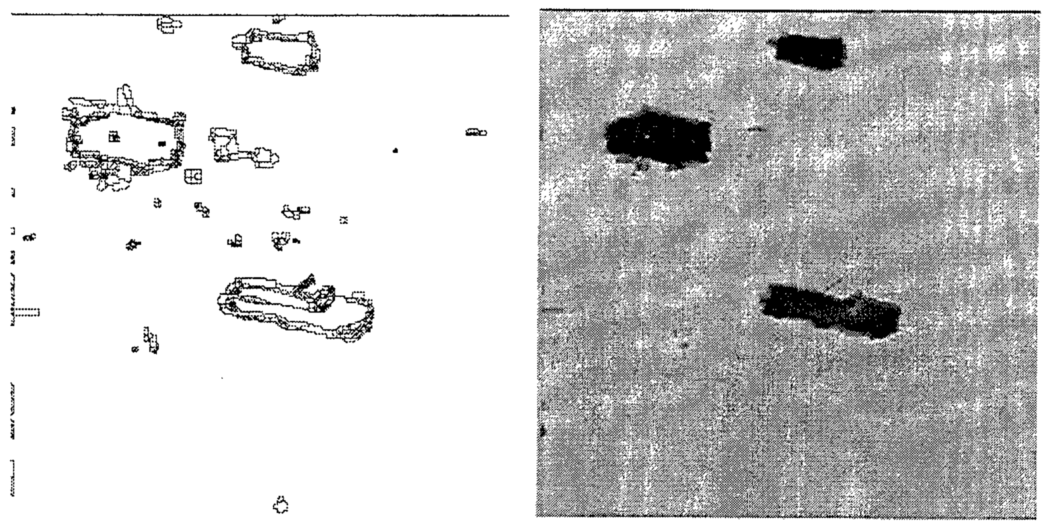


(c) Boundaries from multiresolution segmented image; $\lambda = 32$.



(d) Reconstructed image after multiresolution segmentation; $\lambda = 32$.

FIGURE 6. Three Trucks in Desert Terrain.



(a) Boundaries from multiresolution segmented image; $\lambda = 64$.

(b) Reconstructed image after wavelet segmentation and clutter removal.

FIGURE 7. Trucks in Desert Terrain.

Standard methods such as Wiener filters, constrained least squares, and median filters either blur excessively or suffer from spurious oscillations and are not extendible to more complicated models. The main difficulty is the ill-posedness of the inverse blurring process, which tends to create artifacts near edges and boundaries and to confuse texture with noise. Modern PDE-based techniques borrow ideas from shock calculations, moving front calculations, homogenization, and variational problems with discontinuous solutions. These areas involve the computation of functions with discontinuities and textures, and are core issues of restoration.

A very successful restoration technique was developed based on minimizing the quantity total variation or "wiggliness" of u (e.g., see Figure 5):

$$TV(u) = \int_\Omega \sqrt{u_z^2 + u_y^2}$$

subject to constraints on the blur and noise:

$$\int (Au - u_0) = 0 \text{ (noise has mean zero)}$$

$$\int (Au - u_0)^2 = \sigma^2 \text{ (noise has given variance)} \tag{3}$$

The gradient descent method leads to a time-dependent PDE with initial data u_o, the blurry, noisy image, and constraints enforced using a Lagrange multiplier.

The resulting time-dependent PDE

$$u_t = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda A^* (Au - u_0) + \mu A, \tag{4}$$

with t increasing and λ, μ chosen as to enforce Equation 3 has a solution that approaches the restored image as t increases.

There is an interesting geometric interpretation of Equation 4, linked to mathematical morphology, namely the level sets (contour lines) of u_t move with normal velocity, that equals their curvature, divided by the norm of the gradient of u. Thus, edges (gradient $u = \infty$) don't move, while noise is killed. This is all done while satisfying the constraints. The process can also be interpreted as anisotropic diffusion, where diffusion takes place tangent to, but not normal to, edges—so edges are preserved while noise diffuses away.

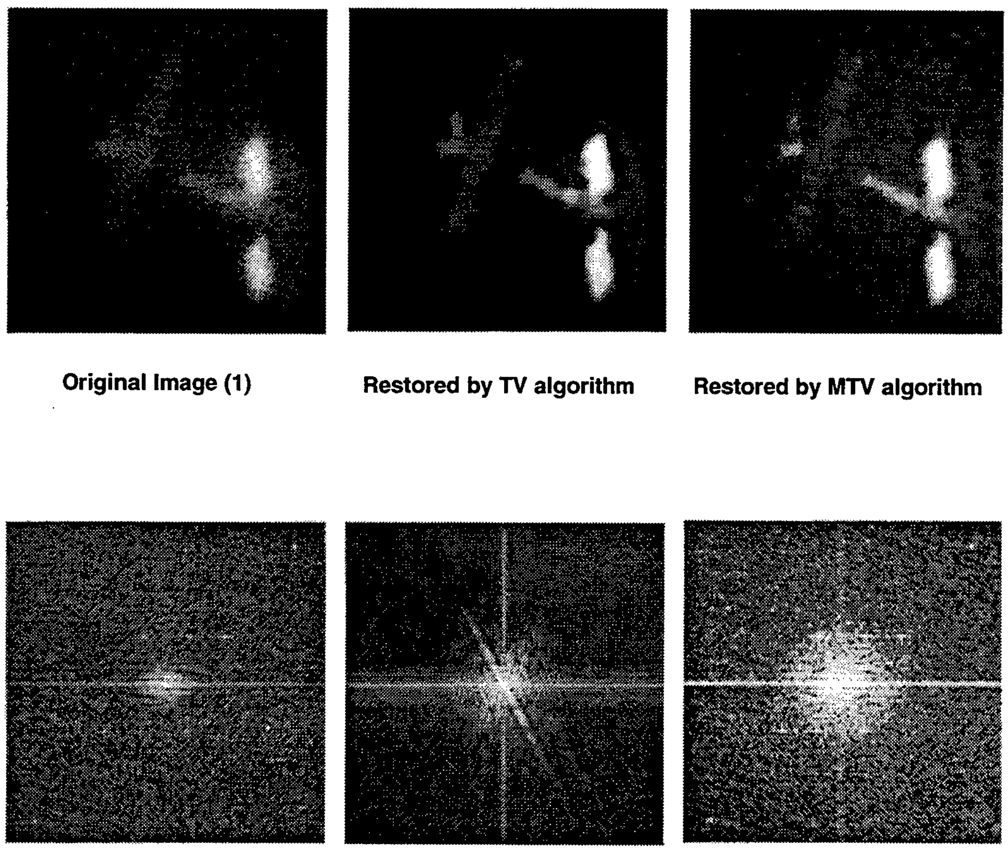
In Figure 8, the left image in the first row is the original satellite image; the middle image has been restored by TV. The right image in the first row shows the exciting possibilities when the restoration is combined with modern multiscale segmentation (MTV) techniques.. The point spread function (e.g., the operator A) is only approximately known, yet these restorations are progressively quite accurate, as the Fourier transforms in the bottom row illustrate. This restoration appears to be state-of-the-art, but improvements are possible, and desired.

Color Processing

Processing of vector-valued image data has obvious applications to color images, where the data are separated into three channels—red, green, and blue or cyan, magenta, and yellow. Other applications include multispectral images, LANDSAT data, multiply sampled images, possibly sampled using different imaging techniques. The PDE methods for image processing developed in France are based on an axiomatic approach of the theory of multiscale analysis of images. These geometrical and morphological axioms can be extended to color images, obtaining systems of equations.

Clearly this area needs to be investigated further so that the processing tasks described and proposed in the previous sections can be extended to color. Preliminary experiments with vector-valued total variation by the University of California at Los Angeles (UCLA) investigator Professor Chan and his students look promising. It does not suffice to separate the color channels and process them separately. These experiments have extended the TV approach to image restoration to vector-valued data in a systematic way. Their "color" TV method automatically adapts to the strength of the individual channels and does not smooth out color edges. Moreover, this method is the correct one in a large framework of possible extensions of the TV approach to vector-valued data. Figure 9 provides sample restoration. Figure 10 shows a serious distortion introduced by channel-to-channel TV (denoising) versus much better result from "color" TV.

SeaSat Image Restoration (1)



Corresponding Log magnitude Fourier Transforms of the top row

Cognitech, Inc.

FIGURE 8. Sea Satellite Restoration.

FIGURE 9. Original Image (upper left), Noisy Tiger Image (upper right), and Restored Image (bottom center).

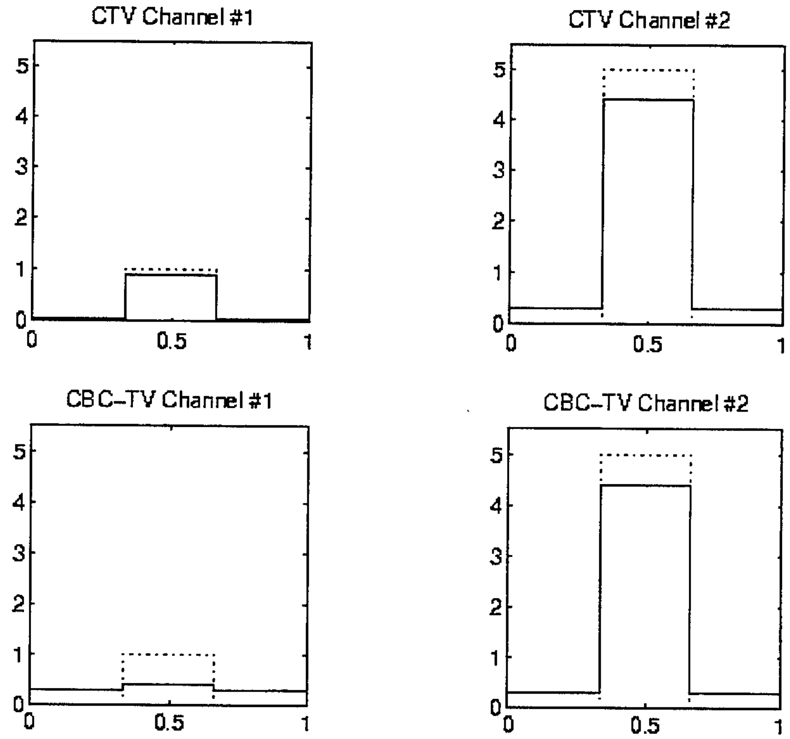


FIGURE 10. The Weaker Channel is Almost Completely Wiped out for Channel-by-Channel Processing, Whereas the Color-TV Processing Automatically Adapts to the Channel Strength.

Blind Deconvolution

Professor Chan using Equation 2 as a model has also been investigating blind deconvolution using TV regularization for both the image and the blur. He has produced very impressive experimental results for realistic images. There is no need to know the class of blur, which itself can have edges (e.g., out-of-focus blur). These results seem much better than the current state of the art. He uses a variational framework in which both the image u and the blur function A are unknown, specifically

$$\min_{u,A} f(u,A) = \frac{1}{2} \|A * u - u_o\|^2 + \alpha_1 TV(u) + \alpha_2 TV(A)$$

The parameter α_1 is related to the signal-to-noise ratio (SNR), whereas α_2 controls the amount of deblurring desired. The current computational algorithm uses an alternating minimization iteration, where each half-step minimizes either with respect to u or A. A continuation method is used in which α_2 is increased from an initial small value, corresponding to the delta-function as blurring operator. Clearly in Figure 11 we can see the object come into focus as α_2 increases up to a critical value; it then becomes defocused again past this value.

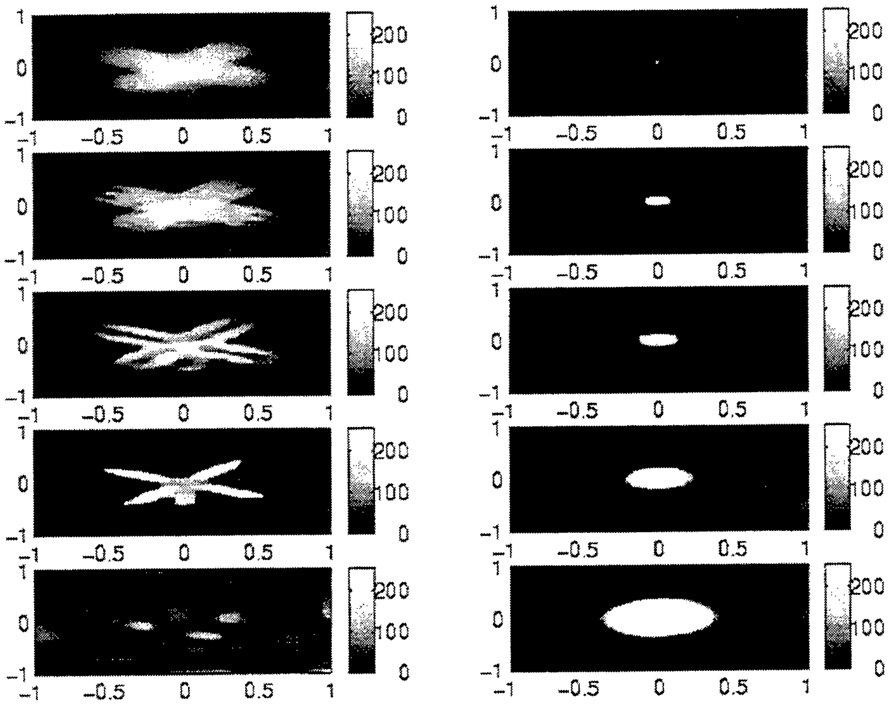


FIGURE 11. Recovered (left) and Identified PSF (right). $\alpha_1 = 2 \times 10^6$; $\alpha_2 = 0, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}$ (from top to bottom).

Image Contrast Enhancement

Images are captured at low contrast in a number of different scenarios. In visible imagery the main reason for this is poor lightning conditions (e.g., pictures taken at night or against the sun rays), while for infrared imagery objects occur at similar temperatures. The gray levels in the digital image may be skewed away from the average luminance, making the detail in the skewed regions imperceptible. As a result, the image is too dark or too bright and is inappropriate for visual inspection or simple observation. The most common way to improve the contrast of an image is to modify its pixel value distribution. A schematic example of the contrast enhancement problem and its solution via histogram modification is shown in Figure 12. On the left, we see a low-contrast image with two different squares, one inside the other, and its corresponding histogram. We can observe that the image has low contrast, and the two objects cannot be identified because the two regions have almost identical gray values. On the right we see what happens when we modify the histogram in such a way that the gray values corresponding to the two regions are separated. The contrast is improved immediately.

As Figure 12 illustrates, the gray-level histogram of an image is a useful representation. For discrete images it is a bar graph whose height represents the frequency of occurrence of the gray-level range in the image, while the different bars correspond to the distinct gray levels. When an image histogram is transformed so that all gray levels occur equally often, the result is an image with a uniform histogram. The rescaled histogram is "flattened" or

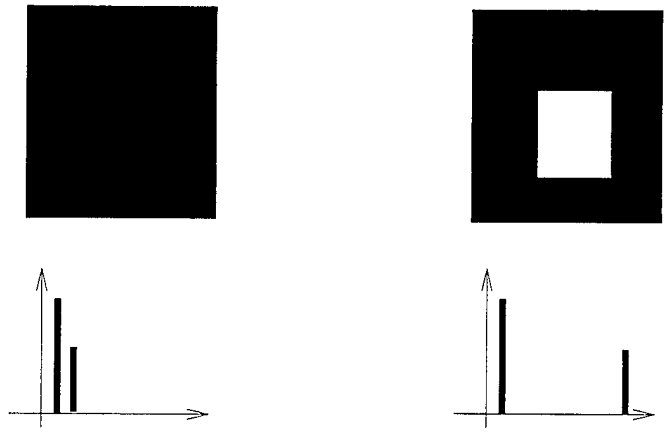


FIGURE 12. Schematic Explanation of the Use of Histogram Modification To Improve Contrast.

"equalized." Equalizing the histogram forces points in the densely populated gray-scale regions (i.e., high amplitude bar graph) to occupy a larger number of gray levels, thus stretching their gray-scale regions. Gray-scale regions in the sparse region are compressed because their points are forced into fewer levels. When the stretched regions are more populous than the compressed regions, the contrast is enhanced. Histogram equalization often performs best on images with detail hidden in the skewed regions. Because there are no limits on how much the image can change, simple equalization does not always improve the appearance, especially with good-quality images.

Although fundamental in image processing, the problem of contrast enhancement still is not solved. Histogram equalization, whose properties are derived by elementary methods in introductory digital-image-processing books, is an unquestioned contrast enhancement technique. The research by Sapiro demonstrates that its subtleties and limitations have not really been revealed by these elementary derivations and that PDE insight can improve even the simplest image-processing technique. He developed image flows and variational approaches as a general framework for contrast enhancement via histogram modification. To equalize the histogram of the initial image $g(x,y)$ we look for solution of the PDE

$$\frac{\partial \Phi(t,x,y)}{\partial t} = \mathrm{A}[(w,z):\ \Phi(\mathrm{t},\mathrm{w},\mathrm{z}) \geq \Phi(\mathrm{t},\mathrm{x},\mathrm{y})]\ -\ \frac{\mathrm{N}^2}{\mathrm{b-a}}(\Phi(t,x,y)-a) \tag{5}$$

where $\mathrm{A}[\cdot]$represents area (or number of pixels in the discrete image $g(x,y)$) and $\Phi_o(0,x,y) = g(x,y)$ is the initial condition. The parameters N, a, and b are related to the gray-scale values in the image. Sapiro shows that the unique steady-state solution of this equation is an image with a uniform histogram. Moreover, the technique extends to image flows that achieve any given pixel value distribution in the steady state.

The following Lyapunov framework shows for the first time histogram modification schemes as image deformations and energy minimization and not just as a gray-value distribution mapping. This framework not only gives new solutions and interpretations to the problem of contrast enhancement, but also opens the door for new algorithms.

The Lyapunov functional for Equation 5, when evaluated along its solutions is decreasing

$$\Lambda(\Phi) = \frac{1}{2}\int\left(\Phi(x,y) - \frac{1}{2}\right)^2 - \frac{1}{4}\iint\left|\Phi(x,y) - \Phi(w,z)\right| \tag{6}$$

Another extension of this technique is to perform local contrast enhancement. Local contrast enhancement is used mainly for object-detection applications. Although the image contrast is significantly improved when performing local histogram equalization, objects are created; see the left panel in the second row of Figure 13. Whereas the traditional histogram approach adds artifacts, the insight garnered from the Lyapunov framework provides an artifact-free solution. The basic idea is to compute the integrals in Equation 6 at specific surrounding areas of the current pixel being modified. The areas of integration are based both on the spatial and gray-value relation between the current pixel and its surrounding neighbors. This technique not only improves the contrast but also preserves shape, as shown in the right panel in the second row in Figure 13. Moreover, it is a major advance in relating contrast changes to image content by understanding the mathematics of contrast enhancement.

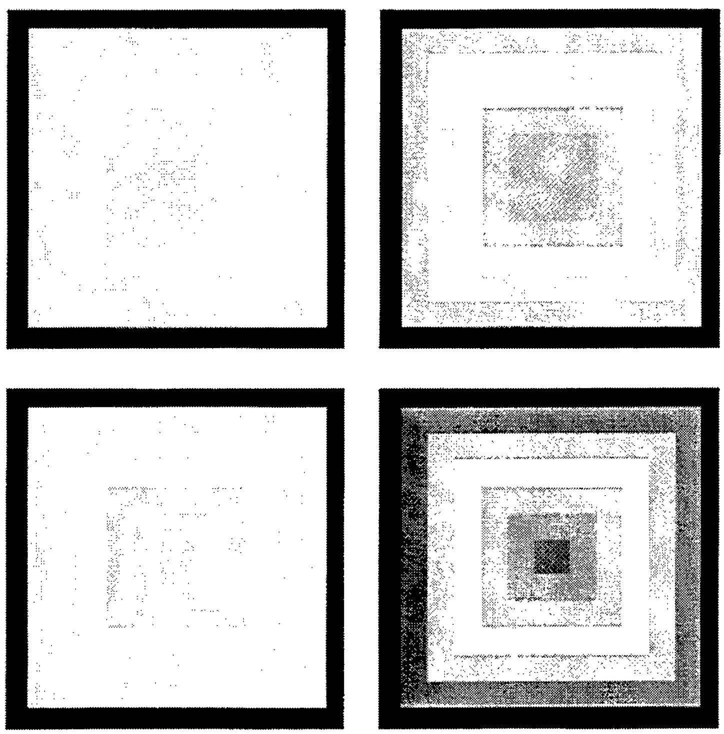


FIGURE 13. Example of Shape-Preserving Contrast Enhancement. Upper left: Original low-contrast image. Upper right: Contrast enhancement via global histogram equalization. Note that—especially for the small rectangles—the contrast is very poor. Lower left: Local contrast enhancement with classical techniques in textbooks. Although the contrast improves, new structures or objects are created (note the new lines). The creation of new structure makes the use of this kind of approach dangerous for detection applications. Lower right: Novel technique for local contrast enhancement (histogram modification) with connected components. Note that the contrast is significantly improved and the structure is preserved.

Image Content Retrieval

Image-content-based retrieval is an emerging research area with application to digital libraries and multimedia databases. As the different sensors are incorporated into the sensor-to-shooter links, the need to store these images in an efficient library integrated with image-exploitation aids will be compelling. While manual annotations can be used to help image search, the feasibility of such an approach to large databases is questionable. In some cases, such as face or texture patterns, simple textual descriptions can be ambiguous and often inadequate for database search. This topic is not technically within our purview, but is in the next step after image enhancement in the process of semiautomatic target generation and assessment of the sort that can dramatically shorten mission planning times. Classification is so sensitive to image enhancement that we did not feel it wise to pursue one without the other. Fortunately, a dramatic new technology has presented itself: wavelet dictionaries, which show promise of changing the face of the subject. This technology allows us to represent many types of signals with very few parameters while retaining key features; this reduces the dimensionality of the classification problem and facilitates its solution.

Image segmentation is one of the fundamental problems in image analysis. Over the past two decades much effort has been spent on detecting edges and image boundaries at multiple scales. Consider, for example, the two-texture mosaic shown in Figure 14 on the left. At the fine scale, one can detect intensity edges making up the texture micro patterns. At the next level of texture discontinuity, one can clearly perceive the inner texture box from the background texture (see Figure 14, center).

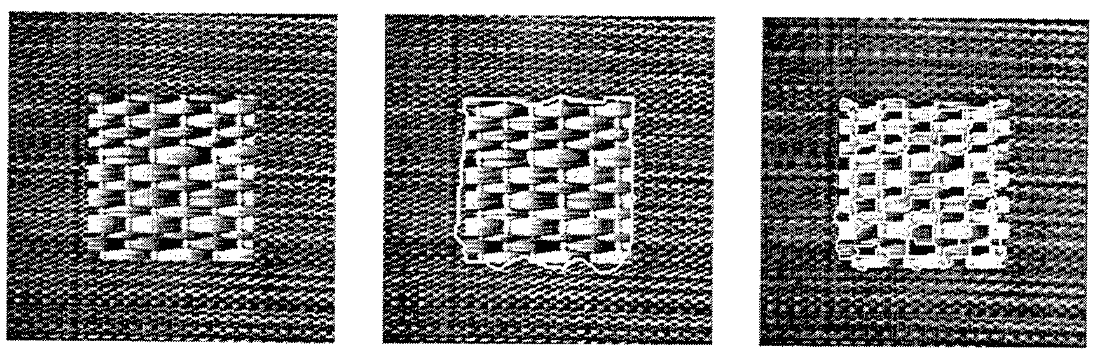


FIGURE 14. Original Image (left), Macro-Segmentation (center), and Micro-Segmentation Within the Central Texture (right).

Professor Manjunath and Dr. Kenney at University of California, Santa Barbara, are developing a framework in which, depending on the scale of interest, one can detect such intensity or texture (and color) discontinuities with very little parameter tuning. Figure 14 (right) shows the results of detecting the micropattern boundaries within the central texture. These texture feature computations are very robust and can be used for pattern retrieval applications. Combining texture and color in computing image boundaries is one of the novel aspects of their approach, and it goes beyond traditional PDE-based image segmentation schemes. In addition, the local image features can be directly used in many pattern retrieval applications, such as ATR and content-based search of large multimedia databases.

Their approach is based on 1) computing the image features, which could be the raw pixel intensity, color, or texture using a Gabor decomposition; 2) computing the local feature gradient; and 3) computing a flow vector at each location based on the gradient energy and direction, and propagating this flow until a stable state is achieved. At the stable state, boundaries are identified as those locations where opposing flow vectors meet or terminate. As shown in Figure 14, the resulting segmentation is very accurate and compares favorably with the state-of-the-art segmentation schemes. Image scale is one of the controlling parameters that directly influences the resolution, accuracy, and number of regions in the segmented image.

Preliminary investigation and implementation of this approach on a large color-image database has yielded very encouraging results, and they are currently integrating this into a content-based retrieval system. This approach has also been used in obtaining a segmentation of airphotos, which is then used for a search based on texture information. A surprisingly diverse set of geographic features (including parking lots, highways, orchards, etc.) can be retrieved using our texture-based indexing of these airphotos, and the method has great promise for automatically maintaining large image/video databases.

Edge Enhancement by Shock Filters

The main goal of edge enhancement is to strengthen edges so that they can be reliably detected; in this view we regard edge enhancement as a preprocessing step to edge detection. Early efforts at edge enhancement used smoothing to eliminate false edges associated with noise. Unfortunately, smoothing such as Gaussian convolution can cause the edges to move or blur. This problem led to traceback schemes in which the edge locations were identified in smoothed images and then followed back through reduced smoothing to their locations in the original image. As might be surmised these schemes were not completely successful.

This situation indicates that an edge-detection preprocessor should satisfy three requirements: 1) the image should be smoothed in such a way that false edges associated with noise or clutter are reduced or eliminated, 2) the true edges in the image should not move, and 3) the true edges should be enhanced.

Edges are the boundaries of image regions. They can be corrupted and obscured by noise, blurring, or low contrast. Low contrast occurs when the jump in intensity across an edge is small. Several successful edge-enhancement schemes, including shock filtering and peer group averaging, use information from the interior of the region to determine the intensity value at the edges. These methods result in enhanced contrast at the edges. At the same time, these methods do not cause unwanted edge motion.

Peer group image processing is related to shock filtering but rises from a completely different paradigm. The shock filtering PDE in one dimension is the convection equation

$$u_t = c u_x \qquad c = -sgn(u_{xx})$$

In peer group processing, each pixel is assigned a peer group, which consists of other nearby pixels with nearly the same intensity value. The average over the peer group is then used to replace the current pixel value. The rationale for this method is that this kind of careful averaging eliminates noise from regions of constant intensity but does not blur or move edges. In fact, peer group averaging can be viewed as a form of shock filtering combined with smoothing perpendicular to the gradient direction. This effect eliminates the problem with salt-and-pepper noise. The number of pixels in the peer group is a parameter that allows one to specify the size of the regions or objects that one wants to selectively enhance.

RESEARCH ABSTRACTS

BROWN UNIVERSITY

Professor Nathan Intrator
nin@cns.brown.edu

Nonlinear feature extraction from high dimensional data representations. This work involves training artificial neural networks based on statistical considerations underlying the projection pursuit framework. Relevant application: nonlinear feature extraction from wavelet and library of basis representations.

Robust classification. This work involves imposing prior knowledge on the classification scheme and interlacing it with the feature-extraction method in such a way that robust classification can be achieved even for moderate size of training data. This work has direct applicability to target and object recognition.

Dr. Quyen Huynh
huynh@cfm.brown.edu

New methods of linear time-frequency (TF) analysis for signal detection. Fixed-subspace TF detection (based on wavelet packet and cosine packet transforms) and variable-subspace TF detection (based on matching pursuit algorithm). These techniques prove to be more robust than matched filter detector and energy detector under mismatch conditions.

Adaptive wavelet analysis of structural acoustic signals. The versatile collection of orthonormal bases allows us to unravel the energy distribution of structural acoustic waves of different scales, such as complex transient dispersive waves.

Design of a robust classifier combining the virtues of modern adaptive time-frequency techniques and BCM optimal selectivity. A network of several BCM neurons, applied simultaneously to different bases of the time-frequency dictionary, encompasses a wide range of time-frequency localization, thus enhancing optimal feature selection. This procedure reduces further the dimensionality provided by the "best discriminating basis."

Advantages of a complete system approach to object recognition. The process of image recognition involves three essential steps: image enhancement, feature extraction, and classification. These tasks are interlaced in such a way that optimal performance can be achieved only if all the blocks are optimized concurrently. In particular, certain types of data deformations and distortion are best addressed at the image-representation level and can be corrected by the image-enhancement scheme, whereas others are better addressed at the feature extraction and classification parts.

As our ultimate goal is optimal automatic classification, it is important to concentrate on those image enhancement features that lead to best classification results. Intrator has introduced a robust method for training neural network classifiers that reduce the sensitivity to various image problems such as blurring, partial occlusion, reduced resolution, and reconstruction from lossy compression. He has found that image blur poses a greater challenge to the classifier than any of the other image deformations and distortions. This suggests a natural combination between image deblurring schemes and robust feature extraction/classification method.

For example, our work has shown that it is possible to train a classifier in such a way that noise, partial occlusion, and certain types of lossy compression have little effect on the performance of the classifier, whereas image blur degrades classification performance severely. Training such a classifier involves optimization of the feature-extraction and classification scheme concurrently.

Nonlinear feature extraction and dimensionality reduction. The tremendous advancement in time/frequency representation of signals has made wavelet dictionaries a natural candidate for preprocessing of complex data. When classification is desired, one has to devise methods for reducing the detailed signal representation to something more manageable by the classifier. Currently, methods for doing that are based on the energy in each time/frequency bin or on linear discriminant analysis.

Huynh and his colleagues at Brown University have introduced a nonlinear feature extraction from the wavelet representations, which shows great promise in classification of transient underwater acoustic signals, e.g., the ability to distinguish between biological clicks and metallic bangs.

Huynh and his colleagues at NUWC apply wavelet analysis to data compression for communication of side scan sonar images. The compression techniques are needed to satisfy real-time processing requirements to transmit these images from remotely deployed MCM vehicles to the host platforms through a single communication channel. We have designed a search algorithm through the wavelet packet trees to find the basis that best characterizes the spatial and spectral characteristics of the underwater mines while maintaining a high compression ratio. The reconstructed images show minimal degradation and still preserve crucial features of underwater mines, which lead to good classification results.

Huynh and his colleagues at NSWC (Carderock Division) have also applied the "best basis" paradigm of Coifman and Wickerhauser to acoustic scattering from a complex underwater vehicle. The results show strong localization in time and frequency, which makes possible the separation of different wave type components of the scattered waves. This type of information is crucial for accurate target detection and identification.

HEWLETT-PACKARD LABORATORIES

Dr. Guillermo Sapiro
guille@hpl.hp.com

Geometric and Invariant Object Detection. We recently developed a stable algorithm for topology independent object detection (with V. Caselles and R. Kimmel). The algorithm is based on geometric PDEs deforming curves toward the boundaries of the objects to be detected in the given image. The algorithm connects, extends, and improves on previous works (e.g., seminal works by Sethian and others), presenting state-of-the-art segmentation results. A number of theoretical results as existence of solutions were proved as well. Since most images are obtained with unknown camera positions, we have extended this theory and developed the first affine invariant geometric active contours for object detection (with P. Olver and A. Tannenbaum). This allows us to perform object detection independently of the (unknown) relative object/camera position.

Vector-valued active contours and anisotropic diffusion. In many applications, images have more than one component. Examples are color images and multispectral data. Multicomponent data can also be obtained from a single image via wavelets type decompositions. This type of decomposition is fundamental, for example, for the segmentation of textures or very noisy images. Based on Riemannian geometry, I have extended the PDE-based segmentation results mentioned above to deal with multicomponent images, presenting the first geometric, topology-free, active contours for the segmentation of vector-valued images. Based on related (Riemannian geometry) theory, we have recently developed algorithms for anisotropic diffusion (enhancement) of vector-valued images (with D. Ringach). Previous geometric approaches mainly performed anisotropic diffusion in each one of the image components independently, without using the high correlation existent between them. This technique can be applied, for example, to color or multispectral images, as well as vector data obtained from wavelets type decompositions.

Shape-preserving contrast enhancement. Most of the work in the area of PDEs in image processing was done for image denoising or deblurring. In computer vision and pattern recognition, the basic work is in shape segmentation and representation. We investigate the use of ODEs and PDEs for one of the most fundamental problems in image processing: contrast enhancement. Images are captured at low contrast in a number of different scenarios. The main reason for this low contrast is poor lighting conditions. As a result, the image is too dark or too bright and is inappropriate for visual inspection or simple observation. The most common way to improve the contrast of an image is to modify its pixel value distribution or histogram. Although fundamental in image processing, the problem of contrast enhancement still is not solved. We have developed image flows and variational approaches as a general framework for contrast enhancement via histogram modification. This framework shows for the first time histogram modification schemes as image deformations and energy minimization, and not just as a gray-value distribution mapping. This work has been extended to perform shape preserving local contrast enhancement (with V. Caselles, J-L. Lisani, and J-M. Morel). This is crucial to guarantee that objects are not created or destroyed while improving the image contrast.

PRINCETON UNIVERSITY

Professor Rene Carmona
rcarmona@princeton.edu

Variational methods for image enhancement and applications. The principal investigator, Rene Carmona, and his research associate, Dr. Sifen Zhong, worked jointly on the two problems discussed separately in the following two paragraphs. In both cases a solution was obtained by solving a variational problem. As is usual in image analysis, such a variational problem consists in the minimization of a penalty function involving a first contribution measuring the smoothness (or lack thereof) of the enhanced image and a second contribution whose intent is to penalize the lack of fit to the original image. The reformulation of standard image-processing problems such as denoising, deblurring, enhancement, and even segmentation as variational problems of this type has proven to be extremely successful in many practical applications of great interest to the Navy, and we plan to further our understanding of these problems and to design new global optimization algorithms to solve them.

The use of nonlinear diffusion equations for image enhancement was made popular by the pioneering work of Perona and Malik and was the motto of Cognitech efforts. The time evolution given by the nonisotropic diffusion equation is of the form

$$\frac{\partial u}{\partial t} = div(g(|\nabla g|\nabla g) + \ldots,$$

where the divergence is defined in terms of the operator $\mathrm{div(u)} = \frac{\partial \mathrm{u}}{\partial \mathrm{x}} + \frac{\partial \mathrm{u}}{\partial \mathrm{y}}$

and the missing terms are of lower order. It can be contrasted with the classical diffusion equation driven by the Laplacian. But more than nonisotropic, this parabolic PDE is nonlinear. This is the main source of the technical difficulties associated with the equation. The interpretation of this nonlinear diffusion equation is that minimal (though variable) smoothing is done in the direction of the gradient (i.e., across the features of the image), and maximal smoothing is done in the direction perpendicular to the gradient (i.e., along the features of the image).

The above discussion shows the crucial role played by the gradient of the image, both through its direction and also via its length. Our first work was to show that this strong dependence of the PDE iterative scheme on the estimation at each step of the direction and the length of the gradient can be very misleading. Smoothing in privileged directions is desirable, but it is a very touchy business, and the actual direction of the gradient can very well be the wrong direction to choose for this directional treatment of the images. Noisy images are the obvious examples that come to mind. However, we proved that there are natural *non-noisy* images for which a blind trust in the direction of the gradient can be damaging. We proposed several alternatives while analyzing the characteristics of the examples we used to illustrate the shortcomings of the gradient. The first alternative consists of choosing the direction of the main component of higher-order derivatives (the Hessian is the simplest case which we considered). A second alternative is to use local spectral features of the image (as given for example with local energy of the Gabor transform). These new ways to determine *a direction along the feature* and a *direction across the features* of an image are more computer intensive than the computation of the direction of the gradient and its orthogonal direction, but they do not have the shortcomings mentioned earlier of the gradient approach. The next step is, on the top of the choice of a couple of orthogonal directions, to find in a data-driven way the amount of smoothing to be done in each direction. The PDE approach has such work automatically done by the diffusivity (i.e., speed) in the Perona-Malik parabolic equation, but nothing is naturally doing the job in the case of the new methods we develop. This is one of our current investigations.

Interior point methods for the enhancement of seabottom images and mine detection. Our second contribution to the analysis of the variational approach to image analysis was to remark that many penalty functions used in the practical implementation of the PDE approach could be minimized using interior point methods. The typical example (and presumably also the most convincing one) is provided by the smoothness penalty defined in terms of the total variation norm advocated first by Fatemi, Rudin, and Osher and more recently by Chan and his collaborators. In this case, if the fit to the data can be expressed in terms of linear constraints, the problem reduces to a *classical* linear programming optimization problem. This reduction to a linear program is accomplished by doubling the number of variables to accommodate the absolute values. Moreover, if the constraints are inequality constraints, one has to include extra *slack* variables. In any case, the dimension of the problem becomes prohibitive. For example, the enhancement of a 128 by 128 image (even with a reasonably small number of constraints) leads typically to a program in 300,000 dimensions! For these reasons, the simplex method cannot be used (at least on standard workstations). But because the matrix of the problem is sparse, interior point methods can be used instead.

We implemented these ideas on the sea bottom side scan sonar images of the database made available by Dr. Gerry Doebeck from the NSWC Coastal Systems Station from Panama City. Our goal was to smooth out all the sources of noise and to enhance all the features that could be *mine-like*. To minimize the number of constraints, we preprocessed the images to search for locations where mines could be present and we placed *local constraints* only in these locations. Our goal is now to extract from these cleaned images the feature vectors from which the detection of the mines will be done without having too large a false alarm rate.

Future Research Directions

We plan to develop the idea of global optimization by interior point methods presented above to the fullest of its potential in the deblurring and enhancing of target images such as seabottom mines. In addition, we are currently implementing a new blind deconvolution procedure using merely a linear program (with an interior point method), and we shall pursue this idea by handling the nonlinearity directly without trying to avoid it by an alternate projection algorithm. The application of our ideas to the problem of blind deconvolution was suggested by T. Chan, who presented some of his preliminary results during the UCLA meeting.

UNIVERSITY OF CALIFORNIA, BERKELEY

Professor James Sethian
sethian@math.berkeley.edu

The University of California, Berkeley, has three research areas.

<u>Diffusion-based schemes for image enhancement and noise removal</u> (Dr. Ron Kimmel, Dr. Ravi Malladi, Professor Sethian). We have developed a set of numerical schemes for image enhancement and noise removal for binary, gray-scale, and color images. The key critical features of our schemes is that they are based on switch functions that hierarchically remove noise based on feature size structure rather than analytic parameters. These schemes stop automatically after the desired size of noise is removed, operate only through nearest-neighbor stencil operations, and are extraordinarily fast.

Shape recovery/edge detection (Dr. Ravi Malladi, Professor Sethian). We have invented, designed, and developed a chain of fast algorithms for shape recovery within images; our primary application so far has been medical images. Our techniques allow us to begin from an initial seed point and quickly locate the desired edge in an automatic fashion; a starting guess near the desired result is not required, nor is an initial shape with the same topological structure. We make use of Sethian's fast marching algorithms coupled to adaptive versions of the Osher-Sethian level set techniques and the image denoising schemes described above to find the boundary in essentially 0(N log N) steps, where N is the total number of points in the domain.

Robotic and path navigation (Dr. Ron Kimmel, Professor Sethian). We have developed fast marching methods which provide 0(N log N) methods for solving static Hamilton-Jacobi Methods, one of which is the standard Eikonal equation. These techniques have been applied robotic tracking, providing very quick optimal paths in variable domain, over terrain, around obstacles, and with orientation dependent effects.

Improving the State of the Art

Each of the cases described above offers significant advances over existing techniques. The image denoising schemes are the first to tackle color in a systematic fashion, and to provide schemes with hierarchically based geometric stopping criteria that automatically halt. Among other things, this means that one does not have to guess á priori how long one must or can run the denoising schemes. Additionally, the fact that the operations are nearest-neighbor stencil operations means that the scheme can be put in hardware very efficiently. The shape recovery/edge-detection work is the state of the art in locating edges or objects within image fields; we suspect that it will be of use in locating particular desired objects within scenes. The work on robotic navigation provides the fastest of all possible algorithms for computing optimal paths in relatively small work configuration spaces.

UNIVERSITY OF CALIFORNIA, LOS ANGELES

Professor Tony Chan
chan@math.ucla.edu

Fast numerical methods for nonlinear PDE image processing models. We develop methods to handle the nonlinearity (using primal dual optimization techniques) as well as preconditioning techniques to handle the differential-convolution operators. These techniques, plus others that we are studying, should make these nonlinear PDE techniques fast and robust enough to be competitive with FFT/wavelet based methods and better at feature extraction. This work is fundamental to all Navy applications using nonlinear PDE methods, especially feature extraction and segmentation.

Extension of total variation models to color and other vector-valued images. We extend TV in a systematic way to vector-valued images. We have shown that our color TV method adapts to the strength of the individual channels and does not smooth out color edges. We hope to apply these techniques to multispectral images. We can also use these techniques in segmentation algorithms. Navy applications include feature extraction and segmentation for color and multispectral images.

Blind deconvolution. Our blind deconvolution method, using TV regularization for both the image and the blur, has produced very impressive experimental results for realistic images. There is no need to know the class of blur, which itself can have edges (e.g., out-of-focus blur). These results seem much better than current state of the art. Navy applications include feature extraction and segmentation.

UNIVERSITY OF CALIFORNIA, SANTA BARBARA

Professor Bangalore Manjunath
manj@ece.ucsb.edu

The information technology has evolved rapidly in the last few years, and as more and more data (images, video, speech) are accumulated, tools for efficient access to the information need to be developed. Content-based search is therefore a key component in managing heterogeneous databases in which it is difficult to use keywords to annotate the information. In managing image/video data, segmentation is the key first step, and our recent efforts in this direction have yielded extremely promising results. We are in the process of developing a new computation approach to segmentation that involves scale space analysis using nonorthogonal wavelet-type Gabor filters. With this approach, we expect to develop a new set of algorithms that are efficient (can be easily implemented in analog/digital hardware) and require very little parameter tuning (unlike most other existing methods in vision). Content-based compression is a potential key application where objects of interest can be selectively coded to achieve very high compression. This will be useful in data storage and transmission.

Our content-based search using texture primitives has demonstrated that one can search for high-level concepts (airport tarmacs, buildings, parking lots) using low-level visual cues. We are expanding this to include multispectral and color information.

Image segmentation using peer group averaging (PGA) (Dr. Charles Kenney at NAWC China Lake and ECE Dept. UCSB, working with Professor B. Manjunath at UCSB and Dr. Gary Hewer at NAWCWPNS). By averaging over the peer group for each pixel, this method retains and strengthens edges and at the same time reduces object regions to a uniform value. The number of elements in each peer group is a parameter for this algorithm that allows one to selectively enhance objects of a given size.

Variational segmentation using boundary functions. The use of boundary functions allows one to give a general framework for variational methods that include standard Mumford-Shah and Geman-type functionals as well as a large class of total-variation-like methods. For all these methods the optimal boundary function can be written in terms of the approximation function u; consequently, the descent PDE for the objective functional can be given in terms of u alone. The explicit form of the boundary function leads to a natural edge-detection method.

Optical flow estimation. If a natural aggregate velocity assumption is used for distant targets with small pixel area, the aperture problem for optical flow can be overcome. This procedure is very fast and avoids dynamic occlusion problems that plague other methods.

How the Work, if Successful, Will Improve the State of the Art

Image segmentation using PGA. A natural advantage of the PGA approach is that it works on the image as a set of discrete intensity values. This is in contrast to image-processing methods based on analogy with smoothly varying continuous surfaces. It has been observed that the PGA method combines the best features of shock filtering and essentially nonoscillatory (ENO) filtering, which currently are considered the best of the state-of-the-art image-processing methods.

Variational segmentation using boundary functions. Unifying the many forms of variational methods under one general framework permits fair comparisons with regard to numerical testing. The use of boundary functions also makes it possible to implement variational methods with boundary functions obtained from other methods, such as PGA or shock filtering. The resulting descent PDEs are linear, as opposed to the nonlinear PDEs that result from descent methods for standard variational methods.

Optical flow estimation. The assumption of an aggregate velocity for intermediate to distant targets eliminates ill-conditioning of the velocity field and has the distinct advantage of speed, because the resulting equations have only two unknowns. It is anticipated that combining this approach with PGA techniques and multichannel information will result in a significant improvement over existing optical flow estimation methods.

Impact on Navy Applications

The work described above is directly related to target acquisition and tracking; because of the speed of these methods, they can be implemented in real-time applications.